



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Control of expression of the I factor, a LINE-like transposable element in *Drosophila melanogaster*

Citation for published version:

Udomkit, A, Forbes, S, McLean, C, Arkhipova, I & Finnegan, DJ 1996, 'Control of expression of the I factor, a LINE-like transposable element in *Drosophila melanogaster*', *EMBO Journal*, vol. 15, no. 12, pp. 3174-3181.

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

EMBO Journal

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Control of expression of the *I* factor, a LINE-like transposable element in *Drosophila melanogaster*

A.Udomkit, S.Forbes, C.McLean,
I.Arkipova and D.J.Finnegan¹

Institute of Cell and Molecular Biology, University of Edinburgh,
King's Buildings, Edinburgh EH9 3JR, UK

¹Corresponding author

***I* factors are LINE-like transposable elements in the genome of *Drosophila melanogaster*. They normally transpose infrequently but are activated in the germ-line of female progeny of crosses between males of a strain that contains complete elements, an *I* or inducer strain and females of a strain that does not, an *R* or reactive strain. This causes a phenomenon known as *I*–*R* hybrid dysgenesis. We have previously shown that the *I* factor promoter lies between nucleotides 1 and 30. Here we demonstrate that expression of this promoter is regulated by nucleotides 41–186 of the *I* factor. This sequence can act as an enhancer as it stimulates expression of the *hsp70* promoter in ovaries in the absence of heat-shock. Within this region there is a site that is required for promoter activity and that is recognized by a sequence-specific binding protein. We propose that this protein contributes to the enhancer activity of nucleotides 41–186 and that reduced *I* factor expression in inducer strains is due to titration of this protein or others that interact with it.**

Keywords: *Drosophila*/*I* factor/transposable element

Introduction

Transposable elements make up a substantial proportion of the genomes of most if not all eukaryotes including man, usually comprising 10–15% of the total DNA. They occur as families of dispersed repeat sequences and can be classified according to their structure and presumed mechanism of transposition. There are two main classes, those elements that transpose by reverse transcription of an RNA intermediate and those elements that transpose directly from DNA to DNA (Finnegan, 1989b).

Elements that transpose via an RNA intermediate are of two types, those that resemble retroviruses in having long terminal direct repeats, namely LTRs and open reading frames similar to *gag*, *pol* and, in some cases, *env* and those that have no terminal repeats and end with A-rich sequences at the 3' end of their coding strands. The latter are often referred to as LINE-like elements as the first examples to be detected were mammalian LINE, or L1, elements. They usually have two open reading frames, the second of which appears to encode a reverse transcriptase. The *I* factor of *Drosophila melanogaster* is of this type.

Transposable elements normally move infrequently.

Precise rates are difficult to determine but most transposable elements in *Drosophila* are thought to transpose at a rate of $\sim 10^{-4}$ – 10^{-5} per generation (Charlesworth and Langley, 1989). In several cases transposition is infrequent because it is restricted by host or element-encoded factors. Transposition of *Ty1* elements in yeast increases 10-fold in strains carrying mutations in *RAD6* (Picologlou *et al.*, 1990) and transposition of the retrovirus-like element *gypsy* is increased in strains of *D.melanogaster* mutant in *flamenco* (Pélisson *et al.*, 1994; Prud'homme *et al.*, 1995) indicating involvement of host factors. Transposition of *P* elements, on the other hand, is regulated by a protein encoded by these elements themselves (Misra and Rio, 1990). Regulatory systems like these are likely to be critical for the survival of a species as increased rates of transposition result in reduced viability, reduced fertility and increased frequencies of mutations.

Transposition of some elements has been shown to be subject to a second level of regulation; in this case one that restricts transposition to the germ-line. This may also be evolutionarily advantageous for both the element and its host since transposition events in somatic cells might debilitate the individual in which they occur without increasing the number of copies of the element transmitted to the next generation. Host genes that regulate transposable elements presumably also have roles to play in the normal life of the organism concerned. A detailed knowledge of mechanisms that control transposition should therefore shed light on factors that control the stability of eukaryotic genomes and allow transposable elements to survive within them while at the same time identifying genes that are important for the host.

The *I* factor is one of a number of LINE-like transposable elements in *D.melanogaster*. Strains containing complete and functional *I* factors are known as inducer, *I*, strains while strains containing only incomplete and non-functional elements are known as reactive, *R*, strains. *I* factors are stable in the genome of an inducer strain but are activated in the germ-line of female progeny of crosses between inducer males and reactive females. This results in reduced fertility and increased germ-line mutations producing a syndrome known as *I*–*R* hybrid dysgenesis. This phenomenon is restricted to females since the male progeny of a dysgenic cross are normal and show no increase in *I* factor transposition. The progeny of crosses between reactive males and inducer females appear normal even though *I* factor transposition can be detected in the female progeny of such a cross at a frequency only ~ 5 -fold less than in dysgenic females (Picard, 1976).

Complete *I* factors are 5.4 kb long and contain two open reading frames (Fawcett *et al.*, 1986). The first, ORF1, encodes a protein with a cysteine-rich motif similar to that found in retroviral *gag* genes and we have shown that it encodes a nucleic acid binding protein (A.Dawson,

E.Hartwood, T.Paterson and D.J.Finnegan, manuscript in preparation). The second, ORF2, encodes a protein with some of the motifs characteristic of a retroviral *pol* gene product and preliminary evidence indicates that it encodes a reverse transcriptase (T.Paterson, A.Gabriel, J.D.Boeke and D.J.Finnegan, unpublished data).

I factors transpose by reverse transcription of a full-length RNA intermediate (Jensen and Heidmann, 1991; Péliou et al., 1991) that can only be detected in tissues in which transposition takes place (Chaboissier et al., 1990). Using constructs in which sequences at the beginning of a complete *I* element were fused to a chloramphenicol acetyl transferase (CAT) reporter gene we have shown that the promoter for this RNA lies entirely within the *I* element itself. The 5' boundary of this promoter lies at nucleotide 1, where transcription is initiated, while the 3' boundary is at about nucleotide 30 (McLean et al., 1993). Transcription from this promoter is regulated in two ways. It is reduced in the presence of complete *I* factors so that transposition is not seen in an inducer strain and it is subject to tissue-specific regulation restricting transcription to the female germ-line (Chaboissier et al., 1990). A similar pattern of expression is seen for the CAT gene if it is linked to nucleotides 1–186 of the *I* factor that make up the 5' untranslated region (McLean et al., 1993).

In this paper we demonstrate that an enhancer that stimulates gene expression in ovarian nurse cells lies between nucleotides 41 and 186. We also show that nucleotides 138–157 are recognized by a sequence-specific binding factor present in nuclear extracts of several tissues including ovaries. These bases are required for full promoter activity and for enhanced expression in ovaries, suggesting that they are recognized by a transcription factor. We propose that this factor, either on its own or in conjunction with other proteins, is responsible for tissue-specific control of *I* factor expression. We also suggest that titration of this protein by complete *I* factors is responsible for the reduced activity of *I* factors in inducer strains.

Results

Tissue-specific control of *I* factor expression

We have identified sequences responsible for tissue-specific expression of the *I* factor by measuring CAT activity in ovaries and non-ovarian tissues (carcass) of females from five lines of reactive flies transformed with a *P* element containing the reporter element, 1-186::CAT. This has nucleotides 1–186 of the *I* factor upstream of a CAT gene with no other promoter. Although the absolute levels of CAT activity in ovary and carcass vary from one transformed line to another, presumably due to the effects of adjacent sequences, there is more CAT activity in ovary than carcass in each case with an average increase of ~20-fold (Table I and McLean et al., 1993).

We have mapped sequences responsible for this effect by measuring the CAT activity in ovaries and carcass of flies carrying deletion derivatives of the 1-186::CAT construct. The results (Table I) indicate that increased CAT activity in ovaries versus carcass requires sequences lying between nucleotides 41–186 and is reduced 5-fold if nucleotides 101–186 are deleted. This suggests that

there is an enhancer element active in ovaries that lies between nucleotides 41 and 186 and that at least some of the sequences required for this are between nucleotides 101 and 186.

In order to confirm that nucleotides 41–186 can act as an enhancer we have introduced nucleotides 41–186 upstream of a minimal promoter for the heat shock gene *hsp70* linked to the *Escherichia coli lacZ* gene in a *P* element transformation vector (Figure 1). Ovaries of females from lines transformed with constructs having nucleotides 41–186 inserted in either orientation were stained for β -galactosidase activity. The patterns of staining seen in ovaries from these flies, as well as in ovaries from control flies carrying the *hsp70-lacZ* genes but without any *I* factor insert are shown in Figure 2.

No β -galactosidase staining was seen in ovaries from control flies (Figure 2A) whereas strong staining was seen in the nurse cells of egg chambers at stages 8–10 in flies carrying the reporter gene preceded by nucleotides 41–186 in either orientation (Figure 2B and C). This confirms

Table I. The levels of CAT activity in ovary and carcass of females from a reactive strain transformed with *I* factor promoter fragments controlling expression of the *E.coli* CAT gene

Promoter	Line	CAT activity		
		Ovary	Carcass	O/C
1-186	132	1807	75	24
	137	1180	32	37
	138	1604	282	5.7
	143	2090	99	21
	148	1740	144	12
				20 \pm 5.3
1-100	178	167	332	0.5
	179	749	82	9.1
	180	1008	476	2.1
	183	920	269	3.4
	184	291	74	3.9
				3.8 \pm 1.4
1-40	177	39	249	0.16
	185	181	193	0.94
				0.55 \pm 0.20

Activity is measured in pmol chloramphenicol acetylated per min per mg protein. The last column shows the ratio of CAT activity measured in extracts of ovary versus carcass together with the mean value for this ratio for all transformed lines carrying the same construct.

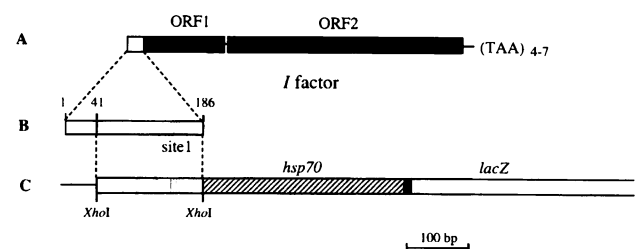


Fig. 1. Map of the *I* factor showing the regulatory region. (A) The structure of a complete *I* factor showing the two open reading frames (shaded boxes) and the 5' UTR (open box). (B) The 5' UTR showing the region, nucleotides 41–186, including the ovarian enhancer element and the transcription factor binding site, site 1. (C) Map of the construct used to demonstrate the enhancer activity associated with nucleotides 41–186. The 282-nucleotide *hsp70* promoter fragment is cross-hatched and the *lacZ* reporter gene is unshaded. The effect of the *I* factor sequence has been tested in both orientations.

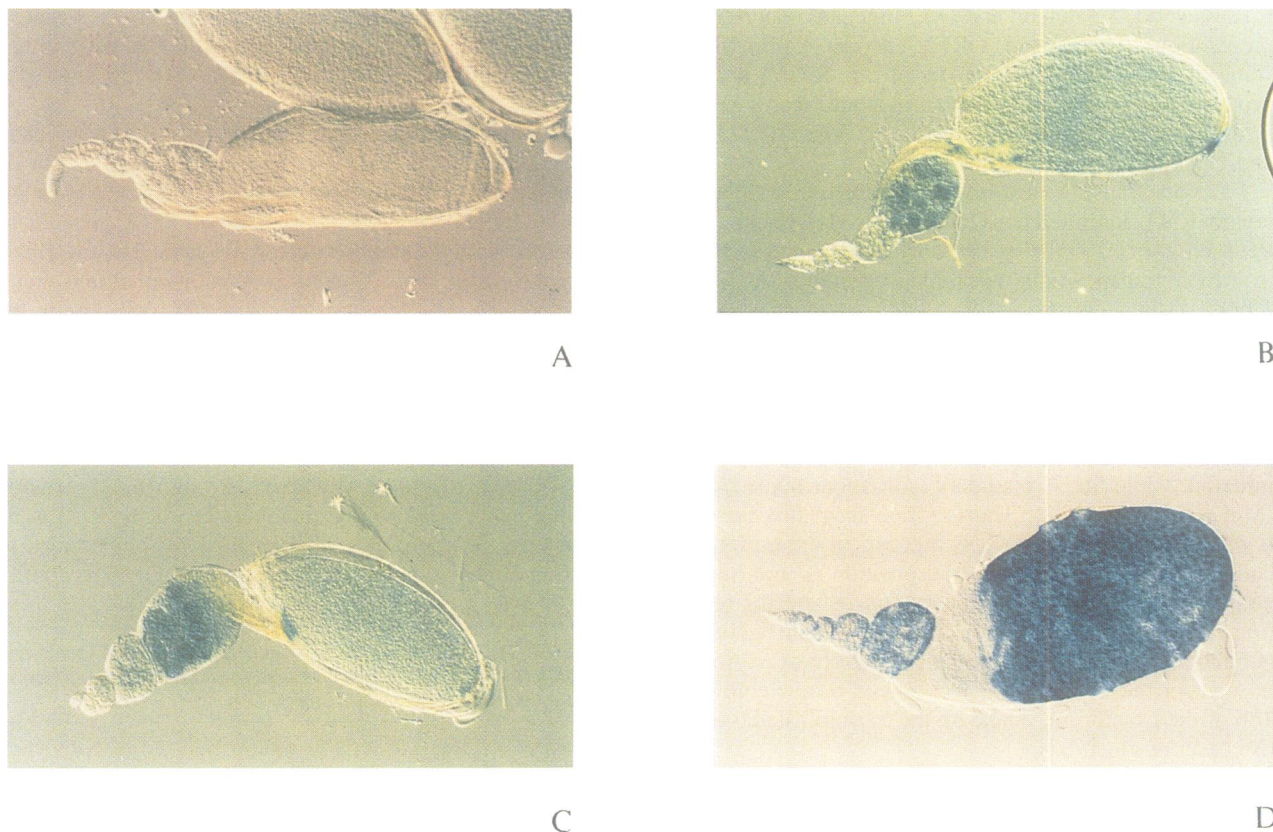


Fig. 2. Nucleotides 41–186 of the *I* factor act as an ovary enhancer. Photomicrographs of ovarioles stained for β -galactosidase and coming from reactive females carrying the *hsp70-lacZ* reporter gene alone (A), or preceded by nucleotides 41–186 in the 5′–3′ orientation (B), or 3′–5′ orientation (C), in the absence of heat shock. (D) The staining of an ovariole from a female carrying the 41-186::*hsp70-lacZ* construct after heat shock.

that there is an enhancer in this region and that it stimulates β -galactosidase expression in female germ cells late in oogenesis. This pattern contrasts with that seen when females from the same lines were heat shocked before their ovaries were stained. In this case β -galactosidase activity was detected in the somatic follicle cells (Figure 2D).

We have estimated the strength of this enhancer by measuring β -galactosidase activity in ovaries of flies from each of two lines carrying these constructs. The level of activity was ~14-fold greater in flies carrying the 41–186 fragment in the 5′–3′ orientation than in flies carrying the *hsp70-lacZ* vector alone (11- and 17-fold in the two lines tested) and ~6-fold greater for flies carrying it in the 3′–5′ orientation (5.4- and 5.8-fold in the lines tested). These results confirm that the effects on CAT expression reported in Table I and by McLean *et al.* (1993) are due to changes in transcription rather than differential RNA stability, a possible interpretation of results from *I*::CAT elements, the transcripts of which contain some *I* factor sequences.

A sequence-specific binding factor recognizes sequences within the enhancer

If nucleotides 41–186 contain sequences that regulate *I* factor expression then these are probably recognized by a transcription factor. We have tested for the presence of such a factor in proteins extracted from the nuclei of various tissues. Figure 3, lanes 5 and 6, shows that a nuclear extract from ovaries of reactive females contains



Fig. 3. Sequence-specific binding of proteins to the *I* factor promoter region. Nucleotides 1–186 labelled with 32 P were mixed with nuclear extracts and competitor DNA before electrophoresis. The extracts used for each lane were as follows: lane 1, none; lanes 2–4, Schneider Line 2 tissue culture cells [SL₂]; lanes 5–7, ovaries of reactive females [OV_(R)]; lanes 8–10, carcasses of reactive females [CAR_(R)]; lanes 11–13, reactive males [♂_(R)]; lanes 14–16, reactive females [♀_(R)]. No competitor DNA was added to the binding reactions in lanes 2, 5, 8, 11 and 14. Cold 1–186 DNA was added to the binding reactions in lanes 3, 6, 9, 12 and 15 (186). Cold 1–Δ186 DNA was added to the binding reactions in lanes 4, 7, 10, 13 and 16 (Δ).

a sequence-specific binding factor that binds to a probe comprising nucleotides 1–186. A preliminary DNase I footprinting experiment suggested that this factor binds to nucleotides 138–157, a region that we refer to as site 1. Further experiments have shown that the protected sequence, ACAAAAACAACAAT, is somewhat smaller, running from nucleotides 137 to 150 (Figure 4). This was the only region that appeared to be protected by this extract.

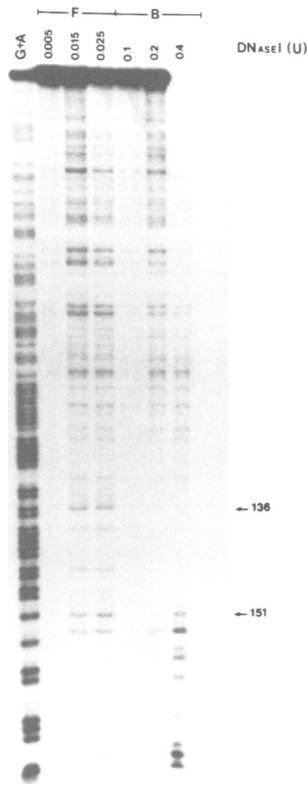


Fig. 4. Identification of site 1 by DNase I footprinting. A DNA fragment comprising nucleotides 1–186 of the *I* factor was labelled with ^{32}P at the 3' end of the top strand. Track 1 shows a G + A sequencing ladder prepared from this fraction. Lanes 2–4 show the result of treating the labelled DNA with 0.005, 0.0015 and 0.025 units of DNase I in the absence of protein. Lanes 5–7 show the result of treating the labelled DNA with 0.1, 0.2 and 0.4 units of DNase I after incubation with 3 µg of an ovarian nuclear extract. The position of the protected sequence has been determined from the G + A sequencing ladder.

We have investigated the occurrence of factors that bind to site 1 in extracts from different tissues using gel retardation experiments (Figure 3). In each case the radioactive probe, nucleotides 1–186, was bound by a factor, or factors, present in nuclear extracts from males (lanes 11 and 12) and females (lanes 14 and 15) from a reactive strain and from *Drosophila* tissue culture cells (lanes 2 and 3). In females this factor was present in extracts of both ovary and carcass (lanes 5, 6, 8 and 9). Similar results were obtained with extracts from individuals of an inducer strain (data not shown).

Site 1 is the target of a sequence-specific binding factor in each case, since nucleotides 1–186 cannot compete for binding if site 1 is not present (Figure 3, lanes 4, 7, 10, 13 and 16) and binding can be competed by a double-stranded oligonucleotide comprising the sequence of site 1 (Figure 5, lanes 5 and 9). We have also shown that there is no detectable binding to a 1–186 probe from which site 1 has been deleted (data not shown).

Three retarded bands, bands 1–3, could be seen when the 1–186 probe was treated with nuclear extracts from males or female carcasses or ovaries (Figure 3). Band 3 was not detected with some extracts of whole females or ovaries but was the only band detected with extracts of tissue-culture nuclei which also yielded some retarded probe at a higher molecular weight. These three bands

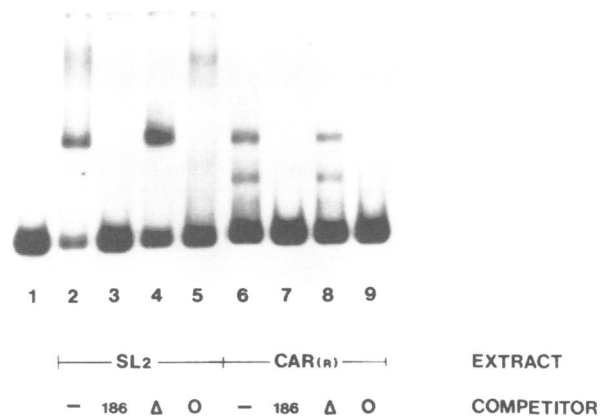


Fig. 5. Nucleotides 138–157 are a target for a sequence-specific binding factor. Nucleotides 1–186 labelled with ^{32}P were mixed with protein extracts and competitor DNA before electrophoresis. The extracts used for each lane were as follows: lane 1, none; lanes 2–5, Schneider Line 2 tissue culture cell [SL₂]; lanes 6–9, carcasses of reactive females [CAR_(R)]. No competitor DNA was added to the binding reactions in lanes 2 and 6. Unlabelled 1–186 DNA was added to the binding reactions in lanes 3 and 7 (186). Unlabelled 1-Δ186 DNA was added to the binding reactions in lanes 4 and 9 (Δ). An unlabelled oligonucleotide comprising nucleotides 138–157 was added as competitor to the binding reactions in lanes 5 and 9 (○).

must be due to the binding of different proteins or different forms or amounts of the same protein. Whichever is the case, these bands are due to factors binding to site 1 since they were eliminated when an excess of intact unlabelled 1–186 DNA was used as a competitor but only if nucleotides 138–157 were present. Differences between extracts may reflect differences in the concentration of binding factor.

Site 1 is part of the *I* factor ovarian enhancer

Since we have shown that nucleotides 101–186 are required for enhanced expression of the *I* factor promoter in ovaries, it is tempting to suggest that site 1 is part of the enhancer. We have investigated whether or not this is the case by measuring CAT activity in ovaries and carcass of reactive females transformed with the CAT gene under the control of nucleotides 1–186 but with site 1 deleted (1-Δ186). Table II shows that deletion of this 20-nucleotide sequence has a dramatic effect, reducing both the absolute level of CAT expression in both tissues and the level in ovaries as compared with carcass. These results indicate that site 1 is indeed required for enhancer activity.

Effect of the inducer state on the *I* factor promoter

We have tried to map sequences responsible for the reduced expression of the *I* factor promoter in inducer strains by making inducer derivatives of each of the reactive strains that carry nucleotides 1–186 or 1–100 linked to the CAT reporter gene. Table III shows the levels of CAT activity in ovaries and carcass in the reactive and inducer backgrounds and that there is a significant reduction in CAT expression in ovaries of inducer as compared with reactive flies with an average difference of ~30-fold. This effect is greatly reduced if nucleotides 101–186 are deleted as the difference for flies with this construct is only ~4-fold. Interestingly the inducer state

Table II. The levels of CAT activity in ovary and carcass of females from a reactive strain transformed with the CAT gene under the control of nucleotides 1–186 but with site 1 (nucleotides 138–157) deleted, together with the ratio of the activity in ovary versus carcass

Promoter	Line	CAT activity		
		Ovary	Carcass	O/C
1-Δ186	9	3.2	9.4	0.34
	11	1.8	5.0	0.36
	13	2.9	7.7	0.38
	14	0.42	2.1	0.2
	17	0.58	0.95	0.61
				0.38 ± 0.07

The mean value for this ratio in the five lines tested is also given.

has little effect on expression of CAT in carcass tissue whichever promoter fragment is used.

These results demonstrate that nucleotides 101–186 not only contain the ovary enhancer but also are responsible for at least part of the inhibition of expression that is seen in an inducer strain.

Discussion

Although complete *I* factors are present in the genomes of inducer strains high levels of *I* factor expression are only seen in the female progeny of a dysgenic cross and then only in their ovaries. This reflects two levels of transcriptional regulation, one that results in higher levels of transcription in ovaries than in female carcass or males, and another that reduces expression in females of an inducer as compared with a reactive strain. We have found that at least some of the sequences responsible for both levels of control are located between nucleotides 41 and 186. Within this region we have identified an enhancer responsible for increased expression in ovaries. This requires site 1 for its activity and we have shown that this sequence is recognized by a sequence-specific binding factor, or factors, present in one form or another in males, female carcass and ovary and *Drosophila* tissue culture cells.

Within ovaries this enhancer stimulates expression in germ-line nurse cells but not in somatic follicle cells. This effect is greatest in egg chambers at stages 8–10. These results are similar to those reported by Lachamme *et al.* (1992) and Tatout *et al.* (1994) who studied β-galactosidase expression in reactive females carrying a translational fusion of ORF1 of the *I* factor (nucleotides 1–290/1014–1104) and the *lacZ* gene. They also detected expression in nurse cells and not follicle cells but could detect β-galactosidase staining as early as stage 2. The reason for this difference is not clear but may be due to the number of copies of the transgene present since Tatout *et al.* (1994) found that they could detect β-galactosidase in the germaria of ovarioles of females with four copies of their construct.

The sequence-specific binding factor, or factors, recognizing site 1 stimulates *I* factor expression since this is reduced if site 1 is deleted (Table II). We presume that this factor affects transcription although we have not demonstrated this directly. Since nurse cells are actively synthesizing host RNA and proteins that are transported

Table III. The levels of CAT activity of isogenic reactive ad inducer females carrying the CAT gene under the control of the *I* factor promoter

Promoter	Line	Ovary		Carcass	
		CAT	R/I	CAT	R/I
1-186	132 ^l	269	6.7	67	1.1
	137 ^l	14	84	17	1.9
	138 ^l	157	10	167	1.7
	143 ^l	99	21	78	1.3
	148 ^l	77	23	129	1.1
			29 ± 14		1.4 ± 0.36
1-100	178 ^l	113	1.5	188	1.8
	179 ^l	261	2.9	65	1.3
	180 ^l	95	10	227	2.1
	183 ^l	325	2.8	180	1.5
	184 ^l	189	1.5	67	1.1
			3.7 ± 1.6		1.6 ± 0.18

Each row indicates the CAT activity for ovary and carcass in females from inducer derivatives of the 1–186 and 1–100 transformants shown in Table I. The ratios of the activity in the reactive versus the inducer background (R/I) are given for each tissue together with the mean values for all transformants carrying the same construct.

to oocytes in preparation for embryogenesis, it is likely that this factor normally functions to stimulate expression of genes coding for some of these products and it may be essential for female fertility. We are trying to identify site 1 binding proteins by molecular and genetic means so that we test this directly.

Although the concentration of site 1 binding protein is similar in all tissues (Figure 3) it only enhances the activity of the *I* factor promoter in ovaries. This may be because it is only in this tissue that it is modified appropriately. Alternatively the activity of site 1 binding protein may require interaction with another protein that is restricted to the female germ-line. If this protein does not itself bind DNA, but interacts with it only through its association with site 1 binding protein, then it would have escaped detection in our DNA binding assays.

The mechanism responsible for tissue-specific expression of the *I* factor contrasts with that utilized by *P* factors, the transposable elements responsible for P–M hybrid dysgenesis, although in each case host-encoded proteins are involved. Transposition of *P* factors is restricted to the germ-line because a splicing factor inhibits removal of the third intron of the transposase gene in the soma (Siebel and Rio, 1990; Siebel *et al.*, 1994) whereas *I* factors exploit a tissue-specific transcription factor.

The activity of another *Drosophila* transposable element, *hobo*, is also restricted to the germ-line because of a transcriptional control (Calvi and Gelbart, 1994). We can detect no similarity between site 1 and the promoter region of *hobo* suggesting that these elements utilize different germ-line-specific transcription factors. This is not surprising as *hobo* is active in both sexes unlike *I* factors that are only active in females.

Although we have been unable to detect any protein binding the *I* factor promoter other than to site 1, we believe that there may be complex interactions regulating *I* factor expression. Deletion of nucleotides 101–186 removes site 1 and reduces expression from the *I* factor promoter in ovaries by ~5-fold (Table I) whereas deletion of site 1 alone has a more extreme effect, lowering

expression in ovaries 100- to 1000-fold and in carcass, 10- to 100-fold (Table II). This may change the spacing of other regulatory factors in this region, cause inappropriate phasing of nucleosomes or allow some residual binding of site 1 binding factor, or of an associated protein so that transcription is inhibited. It is also possible that deletion of site 1 affects translation of *I*::CAT RNA.

The differences between an inducer and a reactive strain are due to the presence of one or more complete *I* factors since a reactive strain can be converted to the inducer state simply by the introduction of a functional *I* factor (Pritchard *et al.*, 1988). When this is done, either by transformation or by a dysgenic cross, the copy number of the *I* factors increases as a result of transposition until it reaches a stable state that must reflect some form of auto-regulation (Pélisson and Bregliano, 1987). The product of ORF2 cannot be involved since females of a reactive strain carrying an *I* factor mutant in this gene do not produce dysgenic progeny when crossed with inducer males (Jensen *et al.*, 1995) indicating that this mutant element can regulate *I* factor expression. This could be mediated by the product of ORF1 as has been suggested previously (Fawcett *et al.*, 1986; Finnegan, 1989a; Busseau *et al.*, 1994), although we have been unable to detect any ability of this protein to bind specifically to nucleotides 1–186 (A.Dawson, E.Hartwood, T.Paterson and D.J. Finnegan, manuscript in preparation), a region that we have found to be sensitive to the presence of complete *I* factors (Table III and McLean *et al.*, 1993). The effect on *I* factor expression of an element with a mutation in ORF1 has not been reported and ORF1 protein could have an effect by interacting with some other DNA sequence or with *I* factor RNA.

If it is not an *I* factor-encoded protein that reduces expression of the *I* factor promoter in inducer strains then maybe it is *I* factor DNA itself. One possible mechanism would be titration of a factor necessary for *I* factor expression. This could be site 1 binding protein itself, as we shall assume in the discussion that follows, or a protein interacting with it.

When a complete *I* factor is introduced into a reactive strain it finds itself, in female germ cells, in a permissive environment with a level of site 1 binding protein that allows transcription and transposition. The resulting increase in the number of copies of site 1 might then titrate this protein until it reaches a point at which transcription falls below the level required for transposition and hybrid dysgenesis.

This mechanism can account for the stimulation of transposition seen in the female progeny of a dysgenic cross since they contain about one half the number of complete *I* factors present in the parental inducer strain. This might reduce the ratio of site 1 DNA to site 1 binding protein sufficiently to allow transposition. This model predicts that we should be able to create a strain that does not permit *I* factor transcription and transposition simply by introducing into a reactive strain multiple copies of site 1 without any coding sequences. We are testing this proposal.

Should these ideas concerning the regulation of *I* factor transposition prove to be correct, then this element is maintained because it takes advantage of host factors both to limit its activity to the germ-line and to maintain its

copy number at a level that is not deleterious to the host but sufficient to ensure its long term survival. This is similar to the mechanism that regulates the expression of *P* elements that are responsible for P–M hybrid dysgenesis since these also utilize a single host factor for both aspects of their regulation. Expression of *P* elements is limited to the germ-line and is only seen in the progeny of a dysgenic cross. In this case tissue-specific control is regulated post-transcriptionally by a tissue-specific splicing event (Laski *et al.*, 1988; Roche *et al.*, 1995) while expression is limited to dysgenic flies because transcription from the *P* element promoter only occurs at high levels in these individuals (Misra and Rio, 1990). These apparently different regulatory mechanisms are linked in that they result from different levels of a particular splicing factor in somatic and germ cells (Siebel *et al.*, 1994; Roche *et al.*, 1995). This factor prevents splicing of the third intron of the transposase gene. If this intron is not removed from *P* factor transcripts then they direct synthesis of a repressor that reduces transcription from the *P* promoter. It is the ratio of the splicing factor to *P* transcripts that determines whether transposase or repressor is produced (Ronseray *et al.*, 1993; Roche *et al.*, 1995).

We believe that the two aspects of *I* factor control are also linked. In this case transcription is restricted to the female germ-line by a tissue-specific transcription factor and is reduced in inducer strains because of changes in the ratio of this factor and its binding site.

Materials and methods

Construction of *I*::CAT reporter elements

P-element transformation vectors carrying fragments of the *I* factor promoter linked to the CAT gene and a polyadenylation signal from SV40 (see McLean *et al.*, 1993 for details) were constructed by ligating *Pst*I fragments containing these sequences to the *Pst*I site of pW8 (Klemenz *et al.*, 1987) so that transcription from the *I* factor promoter would be in the opposite direction to that from both the *white* gene and the promoter of the transformation vector.

Plasmid p1-Δ186 containing the *I* factor promoter with an internal deletion of nucleotides 139–157 was constructed as described by McLean *et al.* (1993) for 3' deletions. The internal deletion of the promoter fragment was synthesized by the PCR using an oligonucleotide primer with an *Xba*I site at its 5' end for the plus strand primer and a primer with a *Bam*HI site at its 5' end for the minus strand. In this case, the minus strand primer was designed so that it contained the sequence complementary to nucleotides 186–128 but with nucleotides 139–157 deleted.

Construction of *I*-lacZ reporter elements

The fragments to be tested for enhancer activity were obtained by PCR using primers with *Xho*I sites at their 5' ends. A fragment containing nucleotides 41–186 of the *I* factor promoter was generated using primers CTCGAGCAGTCTAAAGCCTCGTTTCG for the left-hand end and CTCGAGGATGTGTTGGTTAAGGGCTTG for the right-hand end. This was inserted into the *Xho*I site of pERI vector (Ronaldson and Bownes, 1995) in either orientation just upstream of a minimal *hsp70* promoter upstream of the *E.coli lacZ* gene (Figure 1). This plasmid contains a 282-nucleotide *Xho*I–*Pst*I fragment from the promoter region of the *hsp70* gene inserted upstream of the *lacZ* gene in the *P* transformation vector pCaSpeR-AUG-βgal (Thummel *et al.*, 1988).

P element mediated transformation

The DNA used for transformation of *Drosophila* was prepared using QIAGEN™ columns and was performed by embryo injection as described by Rubin and Spradling (1982). Transposase activity was provided by the helper plasmid phs70Δ2-3wc (McLean *et al.*, 1993) and the recipient embryos were from the reactive strain *W^K* (Luning, 1981). Adults coming from injected embryos were crossed with *W^K* virgins of the opposite sex

and their progeny were examined for a coloured eye phenotype. Homozygous lines were established from any such flies by sibling mating and we have confirmed by Southern blotting that each line used contains a single copy of the transgene.

Inducer derivatives of reactive lines were made by the crossing scheme described by McLean *et al.* (1993).

CAT assays of *Drosophila* tissues

Extracts were prepared from 100 4- to 5-day old male or female flies, 100 pairs of ovaries or 100 female carcass tissues. These tissues were homogenized in 500 µl 0.25 M Tris-HCl pH 7.8. The homogenate was passed through five freeze-thaw cycles and then heated to 65°C for 5 min to inactivate proteins that interfere with the CAT assay. The extract was then spun in a microfuge to pellet cell debris and denatured proteins. The concentration of protein was measured by the Bradford method (Bradford, 1976; Read and Northcote, 1981).

One microgram of extract was mixed with 2 µl 25 mM acetyl coenzyme A, 0.5 µl D-threo-[dichloroacetyl-1-¹⁴C]chloramphenicol (53 mCi/mmol). The volume was made up to 40 µl with water and the reaction incubated at 37°C for 1 h. The reaction was then stopped and extracted by vortexing with 200 µl of ethyl acetate for 1 min. The organic phase was separated by centrifugation in a microfuge for 2 min and was transferred to a new tube and the ethyl acetate allowed to evaporate overnight on bench. Twenty microlitres of fresh ethyl acetate was then added to the tube and vortexed. Acetylated and unacetylated forms of chloramphenicol were separated by silica gel TLC on Merck's TLC silica matrix. The CAT activity was then measured using a PhosphorImager (Molecular Dynamics) and expressed as the number of pmoles of [¹⁴C]chloramphenicol acetylated per min per mg of protein.

Preparation of nuclear extracts from *Drosophila* tissues

The method for preparation of nuclear extracts described by Frank *et al.* (1992) was followed with some modifications. About 500 ovaries from 3- to 4-day old flies were dissected in Ringer's solution. Egg chambers were separated by gently pipetting up and down and were collected by centrifugation at 400 g for 7 min at 4°C. The supernatant was removed and the egg chambers washed in 135 mM NaCl, 5 mM KCl, 1 mM CaCl₂, 0.5 mM MgCl₂, 1.5 mM Na₂HPO₄ and 2.5 mM Tris pH 7.5. The centrifugation was repeated to collect intact egg chambers that were next resuspended in two volumes of 10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl₂, 0.5 mM DTT and 0.5 mM PMSF. The egg chambers were then homogenized thoroughly with a motorized homogenizer and centrifuged at 1400 g for 4 min at 4°C. The pellet was resuspended in an equal volume of 10 mM HEPES pH 7.9, 400 mM NaCl, 1.5 mM MgCl₂, 0.5 mM DTT, 0.5 mM PMSF and 5% glycerol, then 5 M NaCl was added to a final concentration of 0.35 M and incubated on ice for 30 min. The nuclear extract was obtained by high speed centrifugation at 25 500 g overnight at 4°C. Glycerol was added to the supernatant, to a final concentration of 20% and the extract stored in 20 µl aliquots at -70°C. The protein concentration of the extract was determined by the Bradford method. Extracts of other tissues were prepared in a similar way.

Gel retardation assays

Gel retardation was performed as described by Soeller *et al.* (1988). Three micrograms of *Drosophila* nuclear extract were incubated with 1 µg of non-specific competitor DNA [poly(dI-dC), Boehringer Mannheim] on ice in a 10 µl reaction volume containing 25 mM HEPES pH 7.6, 40 mM KCl, 2 mM MgCl₂, 0.1 mM EDTA, 1 mM DTT and 10% glycerol for 10 min. About one nanogram of end-labelled probe DNA was then added to the reaction and the incubation continued on ice for a further 20 min. In competition assays, a 100-fold molar excess of unlabelled DNA was added to the reaction 10 min before the labelled probe. DNA-protein complexes were resolved on 5% polyacrylamide gels in 1× TBE and the position of the probe detected by autoradiography. Extracts were made from the isogenic reactive and inducer strains Charolles and Charolles RC⁺ (Pélisson and Bregliano, 1987) and Schneider Line 2 tissue-culture cells (Schneider, 1972).

DNase I footprinting

DNase I footprinting assays were carried out as described by Leblanc and Moss (1994). The DNA to be protected was a 196 nucleotide PstI-BamHI fragment including nucleotides 1-186 of the *I* factor. This was labelled with ³²P at the 3' end of the top strand using Klenow DNA polymerase. The binding reaction was performed as described for gel retardation. Binding buffer was then added to make a total volume of 50 µl. The reaction tube was transferred to room temperature and 50 µl of cofactor solution (10 mM MgCl₂, 5 mM CaCl₂) was added. Five

microlitres of the appropriate dilution of DNase I were then added. After 2 min of digestion, the reaction was stopped by the addition of 100 µl of stop solution (1% SDS, 200 mM NaCl, 20 mM EDTA pH 8.0 and 40 mg/ml tRNA). The reaction was then extracted with an equal volume of phenol/chloroform (1:1) and DNA precipitated from the aqueous phase with ethanol. The DNA pellet was washed with 70% ethanol, dissolved in 5 µl sequencing loading buffer, denatured by boiling for 2 min and then the fragments separated on a 6% acrylamide sequencing gel. A G + A sequencing ladder prepared from the same fragment as used for the footprint was run in parallel with the sample as a marker.

β-galactosidase staining and enzyme assay

Drosophila tissues were dissected from 2- to 3-day old flies in Ringer's solution. Dissected tissues were left to stain overnight at 37°C in 100 µl of staining buffer solution (10 mM Na₂HPO₄, 150 mM NaCl, 1 mM MgCl₂, 7 mM potassium ferrocyanide, 7 mM potassium ferricyanide and 0.2% X-Gal). The stained ovaries were fixed in 4% formaldehyde in 1× PBS for 15 min and then washed in 1× PBS for 30 min before mounting. Tissues expressing the reporter gene were detected by their blue colour. Heat treatment was performed by placing flies at 37°C for 1 h followed by a recovery period of 2- to 3 h at 18°C before dissection.

For quantitative assays of β-galactosidase activity, ovaries were dissected from 10 flies and were homogenized thoroughly in 500 µl of cold homogenization buffer (60 mM Na₂HPO₄, 40 mM NaH₂PO₄, 10 mM KCl, 1 mM MgSO₄, 50 mM β-mercaptoethanol). Samples were centrifuged at 12 000 r.p.m. for 5 min. Four hundred microlitres of supernatant was pre-incubated at 37°C for 5 min before 600 µl of *o*-nitrophenyl-β-galactopyranoside (ONPG) was added and the incubation continued at 37°C. The enzyme reaction was followed by measuring OD_{420nm} at 10 min time intervals. The protein concentration of each sample was determined by the Bradford method. The β-galactosidase activity of each construct was measured as OD₄₂₀ × 10²/min/mg protein.

Acknowledgements

We are grateful to our colleagues who commented on earlier versions of this manuscript. This work has been supported by grants from the Medical Research Council (G91105876MB), the Wellcome Trust (036248/Z/92/Z/PMG/RB) and the European Union. A.U. is supported by a grant from the Government of Thailand.

References

- Bradford, M.M. (1976) *Anal. Biochem.*, **72**, 248-254.
- Busseau, I., Chaboissier, M.-C., Pélisson, A. and Bucheton, A. (1994) *Genetica*, **93**, 101-116.
- Calvi, B.R. and Gelbart, W.M. (1994) *EMBO J.*, **13**, 1636-1644.
- Chaboissier, M.-C., Busseau, I., Prosser, J., Finnegan, D.J. and Bucheton, A. (1990) *EMBO J.*, **9**, 3557-3563.
- Charlesworth, B. and Langley, C.H. (1989) *Annu. Rev. Genet.*, **23**, 251-287.
- Fawcett, D.H., Lister, C.K., Kellett, E. and Finnegan, D.J. (1986) *Cell*, **47**, 1007-1015.
- Finnegan, D.J. (1989a) In Berg, D.E. and Howe, M.M. (eds), *The I Factor and I-R Hybrid Dysgenesis in Drosophila melanogaster*. American Society for Microbiology, Washington, DC, pp. 503-517.
- Finnegan, D.J. (1989b) *Trends Genet.*, **5**, 103-107.
- Frank, L.H., Cheung, H.K. and Cohen, R.S. (1992) *Development*, **14**, 481-491.
- Jensen, S. and Heidmann, T. (1991) *EMBO J.*, **10**, 1927-1937.
- Jensen, S., Cavarec, L., Gassama, M.-P. and Heidmann, T. (1995) *Mol. Gen. Genet.*, **248**, 381-390.
- Klemenz, R., Weber, U. and Gehring, W.J. (1987) *Nucleic Acids Res.*, **15**, 3947-3957.
- Lachamme, P., Bouhidel, K., Mesure, M. and Pinon, H. (1992) *Development*, **115**, 729-735.
- Laski, F.A., Rio, D.C. and Rubin, G.M. (1988) *Cell*, **44**, 7-19.
- Leblanc, B. and Moss, T. (1994) In Kneal, G.C. (ed.), *DNase I Footprinting*. Humana Press Inc., Totawa, NJ.
- Luning, K.G. (1981) *Hereditas*, **95**, 181-188.
- McLean, C., Bucheton, A. and Finnegan, D.J. (1993) *Mol. Cell. Biol.*, **13**, 1042-1050.
- Misra, S. and Rio, D.C. (1990) *Cell*, **62**, 269-284.
- Pélisson, A. and Bregliano, J.-C. (1987) *Mol. Gen. Genet.*, **207**, 306-313.
- Pélisson, A., Finnegan, D.J. and Bucheton, A. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 4907-4910.

- Pelisson,A., Song,S.U., Prud'homme,N., Smith,P.A., Bucheton,A. and Corces,V.G. (1994) *EMBO J.*, **13**, 4401–4411.
- Picard,G. (1976) *Genetics*, **83**, 107–123.
- Picologlou,S., Brown,N. and Liebman,S.W. (1990) *Mol. Cell. Biol.*, **10**, 1017–1022.
- Pritchard,M.A., Dura,J.-M., Pélisson,A., Bucheton,A. and Finnegan,D.J. (1988) *Mol. Gen. Genet.*, **214**, 533–540.
- Prud'homme,N., Gans,M., Masson,M., Terzian,C. and Bucheton,A. (1995) *Genetics*, **139**, 697–711.
- Read,S.M. and Northcote,D.H. (1981) *Anal. Biochem.*, **116**, 543–578.
- Roche,S.E., Schiff,M. and Rio,D.C. (1995) *Genes Dev.*, **9**, 1278–1288.
- Ronaldson,E. and Bownes,M. (1995) *Genet. Res. Camb.*, **66**, 9–17.
- Ronseray,S., Lemaitre,B. and Coen,D. (1993) *Mol. Gen. Genet.*, **241**, 115–123.
- Rubin,G.M. and Spradling,A.C. (1982) *Science*, **218**, 348–353.
- Schneider,I. (1972) *J. Embryol. Exp. Morphol.*, **27**, 353–367.
- Siebel,C.W. and Rio,D.C. (1990) *Science*, **248**, 1200–1208.
- Siebel,C.W., Kanaar,R. and Rio,D.C. (1994) *Genes Dev.*, **8**, 1713–1725.
- Soeller,W.C., Poole,S.J. and Kornberg,T. (1988) *Genes Dev.*, **2**, 68–81.
- Tatout,C., Docquier,M., Lachaime,P., Mesure,M., Lecher,P. and Pinon,H. (1994) *Int. J. Dev. Biol.*, **38**, 27–33.
- Thummel,C.S., Boulet,A.M. and Lipshitz,H.D. (1988) *Gene*, **74**, 445–456.

Received on November 30, 1995; revised on February 9, 1996